

Global Business Languages

Volume 6 *Proficiency, Efficiency, and Business Languages*

Article 1

May 2010

General and Directed Assessment in Modern Languages

Heidi Rockwood
Georgia Institute of Technology

Follow this and additional works at: <http://docs.lib.purdue.edu/gbl>

Recommended Citation

Rockwood, Heidi (2010) "General and Directed Assessment in Modern Languages," *Global Business Languages*: Vol. 6 , Article 1.
Available at: <http://docs.lib.purdue.edu/gbl/vol6/iss1/1>

Copyright © 2010 by Purdue Research Foundation. Global Business Languages is produced by Purdue CIBER. <http://docs.lib.purdue.edu/gbl>

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Heidi Rockwood
Georgia Institute of Technology

GENERAL AND DIRECTED ASSESSMENT IN MODERN LANGUAGES

As Foreign Language professors, we all like to think that we are of real service to our students; that we teach them a valuable skill and that—in addition to creating a better rounded individual who will have more sensitivity towards cultural differences—we will make them more “marketable,” once they look for a job. Yet the evidence that we can collect to reassure ourselves that we indeed accomplish these goals is generally purely anecdotal. We all have our private success stories, but to what degree do they represent the totality of students who pass through our classes?

Tests may show that our students have mastered certain skills and can react to given signals in artificial settings. They may predict the ability of students to anticipate what certain instructors expect of them, but to what degree are they indicative of real life ability to react to spoken and written commands and signals and to generate the appropriate responses in settings unlike those of a traditional classroom? Unless we develop ways of sending our students into such real life settings and of observing them unobtrusively, we may never know. But the academic community has long tried to come up with more objective and generalized ways of measuring student abilities outside of a particular classroom under the auspices of developing assessment criteria.

In 1990, Catherine Porter Lewis (paraphrasing Peter Ewell) defined the purpose of assessment as “[accumulating] multiple measures (including, but not limited to, existing ‘archival’ data, standardized tests, specially constructed test instruments, and interviews and questionnaires) in order to describe broad outcomes for groups of students...” (Lewis 35). At that time, Lewis (working as part of a team which had the task to develop assessment strategies and instruments for five SUNY campuses) characterized the desirable outcome for foreign language students as a

measurable increase (in terms of proficiency level) in the traditional four skills, in cultural sensitivity and in additional skills such as reading and interpreting literary works. For speaking and writing purposes ACTFL and FSI guidelines and levels served as background for the assessment instruments; tentative measures were developed to check for cultural and literary mastery, also based on the usual pyramid scheme of Novice to Superior skill levels. At the point the article was published, no specific instrument had as yet been devised to measure reading and listening skills (35–39).

The efforts of this particular team are the most comprehensive that I have come across in our own departmental efforts to develop a comprehensive assessment instrument for our institution and its students. About four years ago an official Office of Assessment was founded here with the express purpose of providing information and technical expertise on assessment, of facilitating continuous improvement of academic processes, of disseminating best practices information and of accumulating, generating, and maintaining records based on departmental assessment processes. Interaction between the various schools and departments on campus is fostered through monthly lunch meetings, in which different departmental representatives share their experiences and data with others. Since we are a technological institute and modern languages are at best a fringe interest to students and administrators, however, we have been largely left to our own devices in exploring and testing the assessment instrument we want to use.

The institute by now requires that individual departmental assessment reports be attached to the annual progress report, which is due early in the fall semester. Individual results have to be compressed into brief overviews, which will be shared with other units in the college and serve as basis for the statistics attached to the college's own annual statement. In time, such data are provided to outside reviewers during the regular five year program review.

Since we are a department in which six languages are taught by one to five full-time faculty per language (often a group adds one or more part time faculty members), we decided to leave it up to each group which measures to use for assessment. As a department, however, we decided that we would generally insist only on measures to assess the four skills in the traditional languages and to restrict ourselves more or less to listening and speaking ability in Japanese and Chinese.

In broad terms, all our programs are geared towards teaching languages for professional purposes from the third year onward, with full two semester course series on the third and fourth year level in place in French, German, Japanese and Spanish. In Chinese, so far, courses are available only through the third year (with a business culture oriented set of courses at the third year level and a fourth year under development); and in Russian only first and second year are taught. While literature courses are offered in French, German and Spanish and students are encouraged to take them to enhance their cultural understanding, they do not build the focus of our programs. In fact, the only degree offered by the department is a B.S. in International Affairs and Modern Language (available with tracks in French, German, Japanese and Spanish), for which students take all the courses traditionally required for the B.S. in International Affairs and the institute core, plus a total of 30 hours of their chosen language, built heavily around those above mentioned professional language courses.

It was therefore important to us, to gear all our testing toward a better understanding of business languages rather than keeping it neutral; and the faculty made particular efforts to develop or locate testing instruments which were biased in that direction. Additional measures, such as developing instruments to measure increases in cultural knowledge and understanding or building the above into the tests normally used, was left up to the individual language groups, but those who teach languages for professional purposes were specifically charged with advising others as to the most important skills required in business settings and the best methods of including those skills in assessment efforts and perhaps even in teaching their courses.

After considerable deliberation within the groups themselves and debate within the department, general “desirable outcome” levels were set for the first through the fourth year course levels. The instructors in the oriental languages did not object, even though it seemed initially that the requirements might be too high for their students. (It should perhaps be mentioned in this context that our students come in with an average 1300 SAT and that no degree programs except International Affairs have language requirements. We therefore can count on intelligent, well motivated students who voluntarily attend our courses.) The desirable outcome, or learning objective levels were set as follows:

First Level:

1. Speaking: Students will display the ability to meet the rudimentary requirements of some basic communicative exchanges, such as greeting and leave-taking, describing self and routine activities, ordering meals, asking directions etc.
2. Listening: Students will be able to comprehend face to face and some taped speech, consisting of connected utterances with strong contextual support on familiar, rudimentary topics.
3. Reading: Students will be able to use strategies such as background knowledge, context and cognate clues, to comprehend straightforward, non-complex printed material written for a general audience.
4. Writing: Students will display ability to meet some practical needs such as writing notes, messages, biographical sketches and brief descriptive passages in everyday contexts.

Second level:

- 6) Speaking: Students will display the ability to handle a variety of uncomplicated communicative tasks and social situations with confidence and clarity, including personal description, plans and history and, in general, converse in contexts beyond the most immediate needs with speech that is fairly comprehensible to native speakers. Students will be aware of routine cultural protocols.
- 7) Listening: Students will display the ability to sustain global understanding over longer stretches of connected discourse involving different time frames.
- 8) Reading: Students will display the ability to use effective strategies to negotiate both surface meaning and some cultural inference in authentic literary or periodical texts written for a general readership.
- 9) Writing: Students will display the ability to write for practical needs and limited social demands on familiar topics, and to demonstrate some coherence in description and narration across time frames.

Third/Fourth level:

- (a) Speaking: Students will demonstrate oral interaction and negotiation skills in conflict and professional situations. They will also show historical and cultural awareness in more specialized fields, such as business, technological or literary areas and linguistic analysis.
- (b) Listening: Students will be able to follow native speaker discourse of average speed and complexity.
- (c) Reading: Students will be able to read everything in the foreign language (applies only to the traditional languages) with the help of a dictionary. They will demonstrate some stylistic sensitivity.
- (d) Writing: Students will be able to write about general and some professional topics.

It is understood that the third and fourth level criteria may vary for the different languages and that actual achievement and course content here influence testing. The actual testing instruments chosen by the language groups may therefore differ considerably, especially when it comes to writing skills. Some of the instruments described below are well known, others may not be but are outlined in our Assessment Policy, further explanations follow. The faculty uses all or some of the following instruments:

- OPI (Oral Proficiency Interview). Proficiency levels have been set by ACTFL (American Council for the Teaching of Foreign Languages). The four levels that can be reached by a student during OPI testing (administered one on one by a qualified instructor) reflect considerable internal variety which must be accurately assessed by the tester; however, the general outline is as follows: Novice (none, or only very rudimentary creative use of language patterns and vocabulary); Intermediate (creative, though not necessarily accurate, use of language patterns and vocabulary sufficient for everyday situations); Advanced (creative and mostly accurate use of the language; sufficient for some non-ordinary or conflict situations); Superior (near-native

speaker quality). The OPI assesses speaking and listening ability in the foreign language. Departmental faculty is trained, but not certified as OPI testers. Faculty can assess proficiency levels informally, but do not give out certificates.

- VOCI; a tape-based OPI test, which can be administered to larger groups of students who record their responses on tape. Tapes must be evaluated by qualified and trained instructors. Faculty trained in OPI is also qualified to evaluate VOCI. VOCI tests are available for all the languages taught in the department. VOCI tapes can be ordered from the Language Acquisition Resource Center at San Diego State University; <<http://larcnet.sdsu.edu>>.
- BYU (Brigham Young University) test. A series of computer based, multiple choice questions testing comprehension of written material and grammatical expertise of students. Tests are tailored to each individual student and “designed” by the computer based on ongoing student responses. Point ranges for first year, second year and higher level students are suggested by BYU scoring guidelines, but the latter also point out that it is important that at each individual institutions certain point ranges be established within which the students are expected to perform at a given level. The BYU tests are currently only available for French, German, Russian and Spanish. They are in essence designed to test grammar and vocabulary skills at the lower, and reading ability at the higher levels. To contact Brigham Young University about these tests, call or access the web site. 801.378.4636; <<http://humanities.byu.edu>>.
- Ad hoc testing instruments devised by professional organizations or designed by the faculty of the Department of Modern Languages, such as results of final exams, may be used to assess writing skills.

Once the assessment instruments per se had been decided on, each language group then chose specific ones and determined appropriate scores or test outcomes; choices that were made are outlined below:

Chinese:

Testing instruments are OPI and ACTFL guidelines for reading/writing proficiency.

Scores expected:

- Level 1: OPI Novice High/Intermediate Low; ACTFL rating of Intermediate Low
- Level 2: OPI Intermediate Mid; ACTFL rating of Intermediate High
- Level 3: OPI Advanced; ACTFL rating of Advanced Low

French:

Testing instruments are OPI, BYU and a departmental Prochievement Test.

Minimum satisfactory scores:

- Level 1: OPI Novice High; BYU 200; Prochievement 80
- Level 2: OPI Intermediate Mid; BYU 300; Prochievement 80 (see explanation below)
- Level 3/4: OPI Intermediate High; BYU 400; Prochievement 80

German:

Testing instruments are OPI and BYU.

Scores expected:

- Level 1: OPI Novice High; BYU 300–350
- Level 2: OPI Intermediate Mid; BYU 350–500
- Level 3/4: OPI Intermediate High/Advanced; BYU above 500

Japanese:

Testing Instruments are OPI or VOCI and departmental criteria.

Scores expected:

- Level 1: OPI Novice High; elementary reading/writing
- Level 2: OPI Intermediate Mid; additional reading-writing (approximately 300 kanji)
- Level 3: OPI Intermediate High; 400–500 kanji
- Level 4: OPI Advanced; ability to read the newspaper with the help of a dictionary

Russian:

Testing Instrument is a departmentally designed written proficiency test and VOCI; BYU has only been purchased recently and the scoring ranges for different levels are still being set.

Minimum satisfactory scores:

- Level 1: Novice High/Intermediate Low on the written test and VOCI
- Level 2: Intermediate Mid on the written test and on VOCI

Spanish:

Testing instruments are OPI and BYU.

Scores expected:

- Level 1: Novice High/Intermediate Low; BYU 245–315
- Level 2: OPI Intermediate Mid; BYU 315–375
- Level 3/4: OPI Intermediate High; BYU above 400

After approximately three years of using the above mentioned testing instruments, faculty members are reasonably satisfied with their efficiency in assessing student progress. Most of the criticisms leveled at the assessment instruments were expected. The OPI is time consuming to administer. It is obviously impossible to test all students at a given level but our self-imposed rules have cut down on the sheer numbers involved. We test only students who have taken two courses at a given level from the department and rely more or less on volunteers to recruit participants. Good students generally want to participate; for weaker students point advantages, such as not counting the weakest homework or oral test performance may be given. The decision is up to the individual instructor; all attempt to recruit a representative sample. Students in the business language courses generally participate readily and the OPI is made an integral part of their final exam. The numbers tested are usually larger in the spring semester and while we have not been able to do this at our institution, I would strongly recommend that a course release time be given to an instructor, if he/she shoulders the main burden of OPI testing.

Keeping enough instructors at hand who have OPI training is another, often costly, problem. Since we do not expect faculty to actually become certified testers, the process only involves going through the four day training, which we pay for, but as new faculty is recruited, the expenses for the department are often considerable.

Testing one's own students is another touchy problem in administering the OPI. Obviously instructors tend to play to their students' strengths and the test results are not as objective as one would wish. Some of the time, especially when a new instructor has not yet been trained as tester, different faculty members are asked to administer OPIs, but in essence the faculty candidly admits that the OPI in our setting is more likely to become a prochievement test even at the lower levels, in which the actual achievement expected from a student based on the materials covered in a given course becomes a factor.

French has been the only language group to officially admit that the OPI testing is tainted in this fashion and has set additional clear guidelines for grammar and vocabulary achievement as measured on a 100 point scale (their "Prochievement" test). If students show near-mastery of an expected grammar pattern or vocabulary setting during the OPI, they are also given a certain prochievement rating. Since the general OPI guidelines only establish very few grammar-mastery guidelines (such as question construction or proper use of tenses on the Intermediate level), faculty can be more specific in their prochievement expectations and allot more complete scores. If students are perhaps not very good at using the perfect, but have mastered the case system, their OPI ranking may only be Intermediate Low, but their prochievement score could be quite good, which also adds to the students' feelings of accomplishment. The OPI given after the business language courses has to be prochievement oriented and some cultural and professional skills can be tested by experienced instructors.

On the whole, administering faculty prefer the OPI/Prochievement testing to the other testing instruments, since it assesses two skills—listening and speaking—at the same time and at its best is able to place the student into a near-authentic setting, especially on the business language levels.

The VOPI testing is subject to some of the same objections and has the same advantages as OPI testing. In addition, it is less time-consuming, since a whole class can be tested at the same time and the

instructor does not need to be present. Questions are presented in context by speakers on television and students have to answer by speaking on tape. Obviously, the test cannot be “individualized,” which is both positive, since it is more objective; and negative, since it does not allow for the measuring of achievement levels. The time given for answering a certain question can be limited, so students sitting in close proximity are not tempted to wait for the neighbor’s answer before attempting one of their own. Obviously the time needed for an instructor to listen to all the tapes is considerable, but still less for each individual student than the time needed for an OPI interview, since the instructor only has to listen to the answers. The possibility of playing an answer over again is also appreciated by many.

The real objections leveled by our business language instructors refer of course to the rigid nature of the questions given on television, which are not geared towards specific business situations, even though they may at times cover general cultural skills. More often, questions at the higher levels test historical or social knowledge, which may be peripherally covered in the courses the students have taken, but is not of central concern to them. On those levels, therefore, the OPI is preferred over VOCI. There are plans to create business language versions of VOCI videos, in fact, a Spanish version has been completed, but is currently not available for general purchase, since it is used as the oral portion of the new Business Spanish exam, called EXIGE (Examen Internacional de Negocios in Español).

The BYU test (officially called F- G- R- and S-CAPE for the four languages which use it) is generally the reading test of choice. It is administered via a computer program, which presents multiple choice questions focussing at the lower levels on grammar and vocabulary, on the upper levels on idiomatic expressions and sentence level language samples. It has the obvious advantage that a large number of students can be taken to the computer lab and tested at the same time. Some instructors have even set aside a class period for it. Once the student is in front of the computer, the program reacts to positive and negative results on the first couple of questions rather individually, by either boosting or lowering the difficulty range of the next questions and students in the same class can end up with widely different testing times and difficulty levels. Since the first few responses can be guesses, this type of testing is not always as fair and objective as one would wish and the same student,

taking the test only days or weeks apart, can achieve widely differing results. But measured for a large number of students over several years, the results are fairly steady.

The BYU test is not a testing instrument which is of much use in a business language context. Since the students like it as a measure of general ability, it is also given after two semesters of business language study on the third and fourth year levels, but it hardly reflects specific achievement levels and can at times result in disappointing scores for students whose ability is much higher when they are tested in the appropriate context.

Writing ability is of interest at the upper levels only and individual instructors here have resorted to their own tests to rank students. Obviously all measuring instruments are highly achievement oriented, with some instructors giving tests which explicitly ask students to integrate certain business area vocabulary, grammatical and idiomatic constructions into their tests. Objective standards, as to what can be expected of students at a given level are hard to come by. Some indications can be taken from publications such as *Kursstrukturen*, a sample of syllabi and tests for all levels of business German. San Diego State University at one point decided to put together a similar syllabi and test bank for other languages; but I have not been able to ascertain whether it was or is available.

In allowing a large amount of achievement-directed, individual leeway in testing and evaluating, one might wonder whether one does not defeat the purpose of assessment and whether creating individualized, institute- and course-specific instruments does not eliminate all objective relationship to national standards. But the standards that exist nationally are vague and perhaps even unfair, and the range of achievements which students should be able to produce are generous. The BYU tests come with prepackaged point ranges for given levels that are very wide and will probably fit most any institution and student population, making them of little objective use. On the other hand, using for example FSI standards of oral competence which relate a certain number of weeks or years to a certain level of achievement, often gets institutions into trouble, since they are using fewer teaching hours per week with significant breaks between them and therefore are not getting their students to the predicted achievement levels. Thus some individualization of tests and expected results seems only fair.

Departments might even be surprised at their own successes. Especially at our technologically oriented institution, where we have been able to set up only three hours per week for classes at all levels in most languages (Chinese and Japanese offer four per week in the first year), we originally expected our students to perform at lower than predicted levels. That did not prove to be the case. Achievements have generally been higher and test scores better than predicted and in setting the final Assessment Policy, we had to move up expectation levels a good bit. It might be an interesting research topic to find out whether general motivation (as mentioned above, our students are not taking required courses) and original SAT scores are better predictors of achievement levels than the number of classroom and lab hours per week or the prestige of the study subject at a given institution (it happens to be low at ours).

One of the main thoughts behind assessing student progress is, of course, the idea that it can and will guide teaching techniques and textbook choices especially at the first and second year level to enable students to do well at the upper levels. Since we started some tentative assessment processes in the early nineties, about the same time as most of the upper level business language programs were being set up (only German was fully developed already), I can state with certainty that it has had exactly that effect on us. In retrospect, the first year level seems to be least affected. While all language groups now use proficiency-oriented textbooks, that was pretty much the case before assessment. The choice of second year texts, however, is already more critical. The language groups use texts that focus on general cultural background and only marginally on preparing students for reading and interpreting literary texts. In general conversation and composition courses on the third year level, most instructors now use web-based materials that vary with each time the course is taught and are heavily geared towards professionally useful skills.

As Lewis points out in the 1990 article (38), results of any assessment project are not easily transferable. My hope, however, is that outlining the general process of creating an assessment instrument and directing colleagues to some tests that have proven useful may provide suggestions and perhaps shorten the process of setting up useful procedures at other institutions.

WORKS CITED

- Broschek, Erika, Bettina Cothran and Michael Nentwich, ed.
Kursstrukturen. New York: Goethe Institute, 1995.
- Lewis, Catherine Porter. "Assessing the Foreign Language Major at the
State University of New York: An Interim Report." *ADFL Bulletin*
21.3 (1990): 35–39.